

## ARTICLE

Received 9 Sep 2013 | Accepted 4 Feb 2014 | Published 26 Feb 2014

DOI: 10.1038/ncomms4378

# Surrogate reporter-based enrichment of cells containing RNA-guided Cas9 nuclease-induced mutations

Suresh Ramakrishna<sup>1</sup>, Seung Woo Cho<sup>2</sup>, Sojung Kim<sup>2</sup>, Myungjae Song<sup>1</sup>, Ramu Gopalappa<sup>1</sup>, Jin-Soo Kim<sup>2</sup> & Hyongbum Kim<sup>1</sup>

RNA-guided endonucleases (RGENs), which are based on the clustered, regularly interspaced, short palindromic repeat (CRISPR)-CRISPR-associated (Cas) system, have recently emerged as a simple and efficient tool for genome editing. However, the activities of prepared RGENs are sometimes low, hampering the generation of cells containing RGEN-induced mutations. Here we report efficient methods to enrich cells containing RGEN-induced mutations by using surrogate reporters. HEK293T cells are cotransfected with the reporter plasmid, a plasmid encoding Cas9 and a plasmid encoding crRNA and tracrRNA, and subjected to flow cytometric sorting, magnetic separation or hygromycin selection. The selected cell populations are highly enriched with cells containing RGEN-induced mutations, by a factor of up to 11-fold as compared with the unselected population. The fold enrichment tends to be high when RGEN activity is low. We envision that these reporters will facilitate the use of RGEN in a wide range of biomedical research.

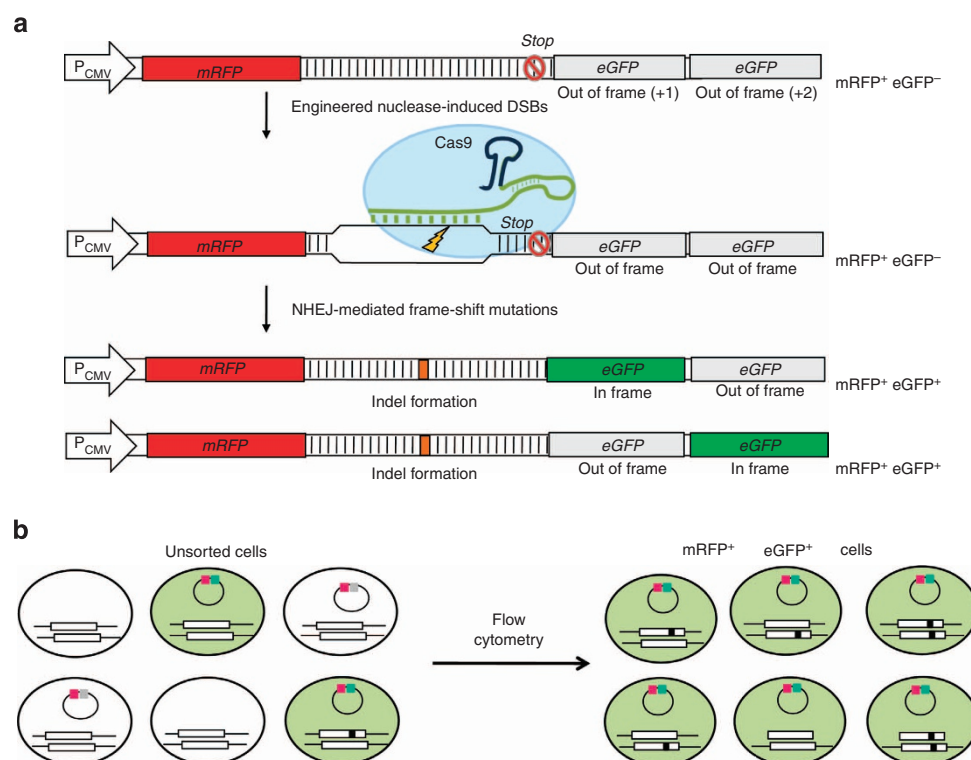
<sup>1</sup>Graduate School of Biomedical Science and Engineering/College of Medicine, Hanyang University, Seongdong-gu, Seoul 133-791, South Korea. <sup>2</sup>National Creative Research Initiatives Center for Genome Engineering and Department of Chemistry, Seoul National University, Gwanak-gu, Seoul 151-747, South Korea. Correspondence and requests for materials should be addressed to H.K. (email: hkim1@hanyang.ac.kr) or to J.-S.K. (email: jskim01@snu.ac.kr).

**R**NA-guided endonucleases (RGENs)—programmable nucleases derived from the RNA-guided CRISPR/Cas DNA cleavage system found in bacteria and archaea<sup>1,2</sup>—are emerging as a promising tool for targeted genome editing<sup>3–10</sup>. As part of their adaptive immune system, bacteria and archaea integrate short fragments (~20 bp) of invading foreign DNAs, termed ‘protospacers’, into their genomes at sites called CRISPRs (clustered regularly interspaced short palindromic repeats)<sup>2</sup>. Three types of CRISPR systems have been identified; type II systems have been utilized for genome editing<sup>3–10</sup>. In such systems, the protospacers in the CRISPR are transcribed to give rise to pre-CRISPR RNA, which is subsequently processed to CRISPR RNA (crRNA) by RNase III. When CRISPR-associated protein 9 (Cas9) is complexed with crRNA and tracrRNA, it forms a sequence-specific endonuclease that recognizes the sequence complementary to that of crRNA and makes double-strand breaks. RGENs can be simplified by replacing crRNA and tracrRNA with a single-chain guide RNA (sgRNA) generated by combining crRNA and tracrRNA<sup>3</sup>. Genome editing using RGENs has been recently reported in various cell types and species such as bacteria<sup>9</sup>, plants<sup>11–13</sup>, *C. elegans*<sup>6</sup>, fruit flies<sup>7</sup>, zebrafish<sup>8</sup>, rats<sup>14,15</sup>, mice<sup>5,15,16</sup> and human cells<sup>4,5,10,17</sup>, including human pluripotent stem cells<sup>10,18,19</sup>.

However, not all designed RGENs exhibit high activity<sup>4–8,10–18</sup>, hampering their use for genome engineering. Thus, simple methods for enriching cells containing RGEN-induced mutations would facilitate the use of this tool. If nuclease-induced mutations confer selective advantages on the cells, these modified cells can be enriched<sup>20</sup>. However, this approach cannot be applied to most genes whose disruption does not contribute to discernible phenotypes for selection. Gene-modified cells are often enriched

by selecting cotransfected cells; for example, transcription activator-like effector nuclease (TALEN)-modified cells have been enriched by selecting transfected cells via flow cytometric sorting or antibiotic selection using vectors expressing fluorescent proteins or antibiotic resistance factors<sup>21–23</sup>. However, even in the transfected cells, nuclease-induced genome editing often occurs only in a limited fraction<sup>24</sup>. Thus, a method to select cells with high programmable nuclease activity would be an efficient approach to enrich cells with nuclease-induced mutations. We and others have previously observed that, when a target sequence on a chromosome was mutated by zinc finger nucleases (ZFNs), the same target sequence on the other homologous chromosome in the same cell is mutated more frequently than that in other cells<sup>25,26</sup>. This observation led us to postulate that an episomal surrogate reporter that contains the target sequence might faithfully reflect the nuclease’s activity on the chromosome in the same cell. This hypothesis was validated as we previously showed that surrogate reporters, together with flow cytometry, magnetic separation or hygromycin selection, can enable enrichment of cells containing mutations induced by two other types of engineered nucleases, ZFNs and TALENs<sup>24,27</sup>.

RGENs differ from ZFNs and TALENs in several key respects. For example, both ZFNs and TALENs (which function as heterodimers) are composed solely of protein, whereas RGENs (which function as monomers) consist of protein and RNA. As a result, ZFNs and TALENs recognize target sequences via DNA–protein interactions, whereas RGENs do so by DNA–RNA interactions. Furthermore, the ZFN and TALEN catalytic domain is derived from the FokI nuclease, whereas the RGEN catalytic domain is from Cas9. Functionally, ZFNs (and probably TALENs) generate double-strand breaks in DNA with cohesive



**Figure 1 | Overview of flow cytometric enrichment of cells containing RGEN-induced mutations. (a)** The working mechanism of the fluorescent reporter. mRFP is constitutively expressed by the CMV promoter ( $P_{CMV}$ ), whereas eGFP is not expressed without RGEN activity because the eGFP sequence is out of frame and there is a stop codon before eGFP. If a double-strand break (DSB) is introduced into the target sequence by an RGEN, the break is repaired by error-prone nonhomologous end joining (NHEJ), which often results in indels. This indel formation can cause frame shifts, making either of the eGFP genes in frame and expressed. **(b)** A schematic illustrating the enrichment of mutant cells in the mRFP<sup>+</sup> eGFP<sup>+</sup> cell population using flow cytometry. Reporter plasmids, chromosomal target loci and sites of mutations, represented by black boxes, are shown.

overhangs, whereas RGENs mainly lead to blunt ends. Given these unique features of RGENs, we examine whether cells containing RGEN-induced mutations can also be enriched with surrogate reporters.

Here we show that three kinds of episomal surrogate reporters, coupled with flow cytometry, magnetic separation or hygromycin selection, can indeed enrich RGEN-induced mutant cells, which should facilitate the use of RGENs in biomedical research, medicine and biotechnology.

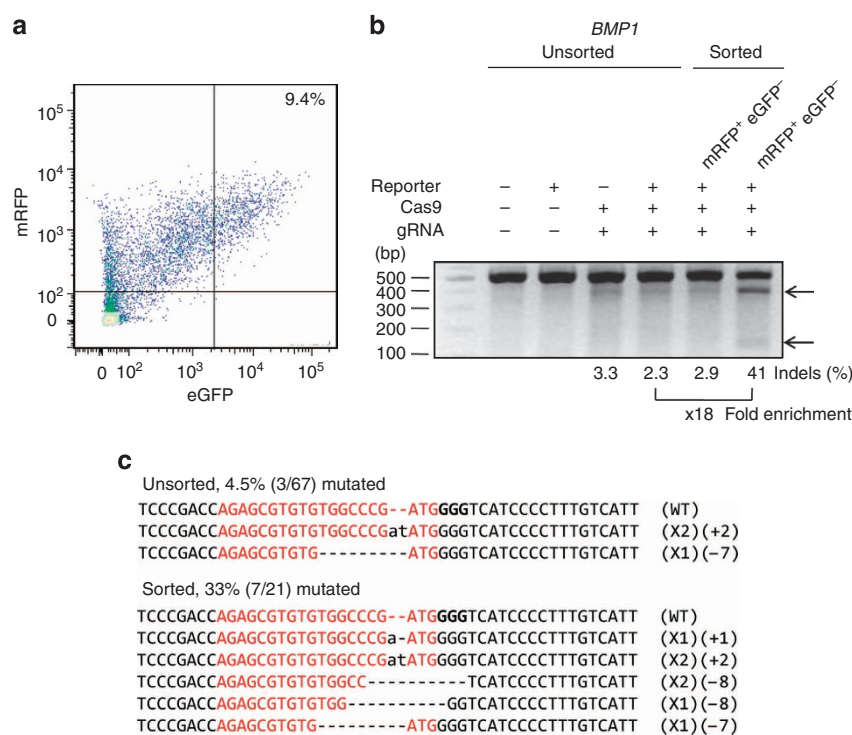
## Results

### Flow cytometric enrichment of cells containing mutations.

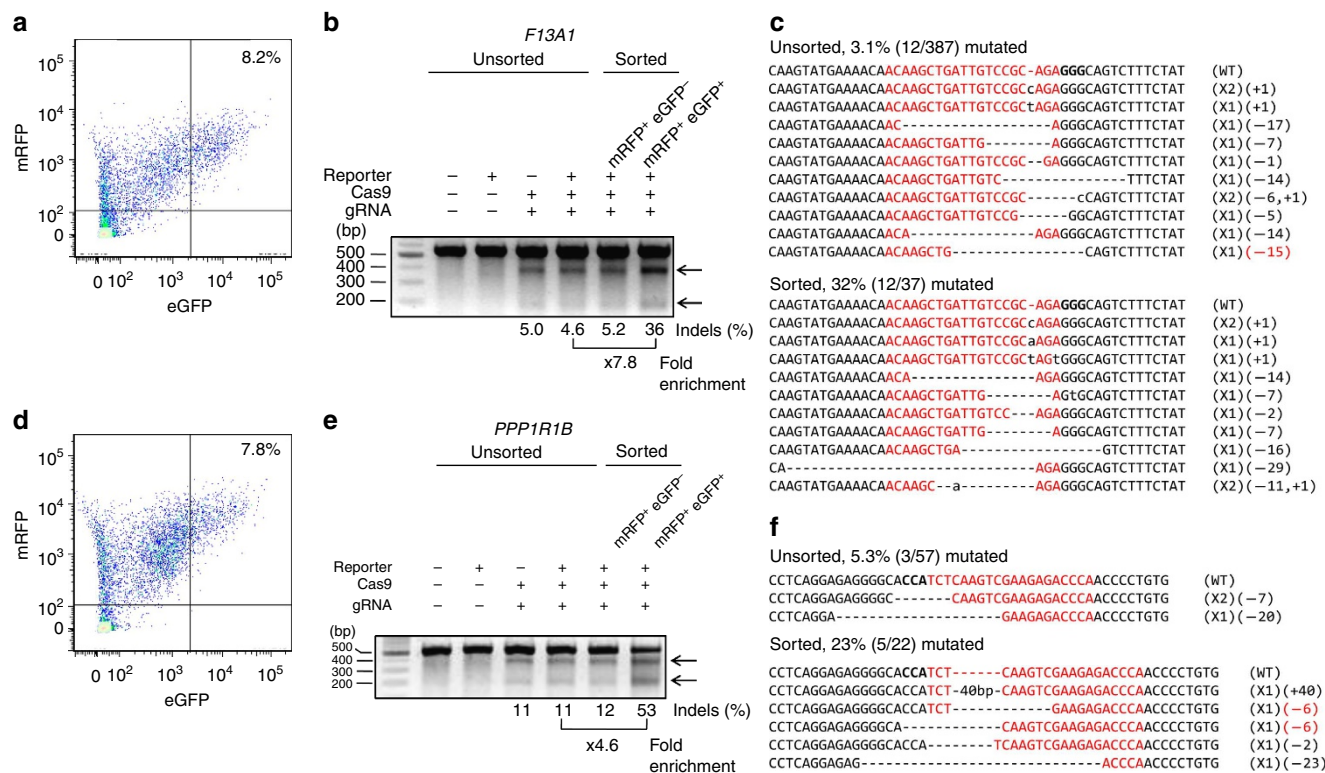
The reporter comprises an mRFP gene, an RGEN target sequence and two eGFP genes (Fig. 1). To increase the sensitivity of the previously described fluorescent reporter<sup>24</sup>, we added another eGFP-encoding sequence at the end of the original eGFP sequence such that mutations of both (3n+1) and (3n+2) nucleotide insertions or deletions at the target sequence can lead to eGFP expression. In the absence of nuclease activity, the reporter-transfected cells will express only mRFP owing to the presence of double barriers: the two eGFP genes are both out of frame and the stop codon before the eGFP genes will terminate protein translation. If RGENs make double-strand breaks in the target sequence in the reporters, indels generated by error-prone nonhomologous end joining (NHEJ) can cause frame-shift mutations, rendering one of the eGFP genes expressed (Fig. 1). To determine whether mutant cells can be enriched using flow cytometric reporters, we cotransfected a reporter plasmid, a plasmid encoding Cas9 and a plasmid encoding BMP1-targeting

crRNA and tracrRNA into HEK293T cells. Twenty-four hours after the transfection, a considerable fraction of cells expressed mRFP, whereas a small fraction of cells expressed eGFP (Supplementary Fig. 1), indicating that frame-shifting indels had already been generated at the target sequence by RGEN activity. In contrast, a previous study involving ZFNs showed that eGFP expression was barely observed at day 1 after transfection<sup>24</sup>, suggesting that RGENs may work faster than ZFNs. This observation is compatible with a recent report that RGEN-mediated gene editing can be detectable at ~20 h post transfection compared with ~40 h for TALENs<sup>10</sup>. The number of eGFP-expressing cells also gradually increased over 3 days, suggesting that RGEN activity was also present after 24 h post transfection.

Three days after transfection, 9.4% of the cells were mRFP<sup>+</sup>eGFP<sup>+</sup> cells; these cells were flow cytometrically separated (Fig. 2a). To measure the RGEN-induced mutation frequencies (or indel %), we isolated genomic DNA and performed a T7E1 endonuclease I (T7E1) assay<sup>26,28</sup>. T7E1 recognizes and cleaves heteroduplexes formed by the hybridization of wild-type and mutant DNA sequences. This assay showed that the mutation frequency at the BMP1 gene in mRFP<sup>+</sup>eGFP<sup>+</sup> cells was 41%, 18-fold higher than that in unsorted cells (2.3%) (Fig. 2b), indicating drastic enrichment of BMP1-mutated cells. Comparable mutation frequencies were obtained in cells transfected with only Cas9 and guide RNAs (3.3%), unsorted cells (2.3%) and mRFP<sup>+</sup>eGFP<sup>-</sup> cells (2.9%), which is in line with previous observations using ZFNs<sup>24</sup>. When we determined the DNA sequences around the target site, we found that the mutation frequency in the flow cytometrically sorted



**Figure 2 | Flow cytometric enrichment of BMP1-disrupted cells.** (a) Flow cytometry of HEK293T cells at 3 days after cotransfection of the reporter plasmid, a plasmid encoding Cas9 and a plasmid encoding BMP1-targeting crRNA and tracrRNA. The percentage of mRFP<sup>+</sup>eGFP<sup>+</sup> cells is shown. (b) The frequency of RGEN-induced mutations as determined by the T7E1 assay. Arrows indicate the expected positions of DNA bands cleaved by mismatch-sensitive T7E1. The numbers at the bottom of the gel indicate mutation percentages calculated by band intensities. (c) DNA sequences of the wild-type (WT) and mutant clones, with RGEN recognition sites shown in red and the protospacer adjacent motif (PAM) sequence in bold characters. Dashes and lower-case letters indicate deleted and inserted bases, respectively (the number of inserted or deleted bases are described in the parentheses; out-of-frame and in-frame mutations are indicated using red and black colours, respectively). The number of occurrences is shown in parentheses; X1 and X2 indicates the number of each clone. Mutation frequencies were obtained by dividing the number of mutant clones by the number of total clones.



**Figure 3 | Flow cytometric enrichment of *F13A1* and *PPP1R1B*-disrupted cells.** Three days after cotransfection of the reporter plasmid, a plasmid encoding Cas9 and a plasmid encoding *F13A1*- (**a-c**) and *PPP1R1B*- (**d-f**) targeting crRNA and tracrRNA, HEK293T cells were flow cytometrically sorted and analysed. (**a,d**) Flow cytometry; the percentage of mRFP<sup>+</sup>eGFP<sup>+</sup> cells is shown. (**b,e**) The frequency of RGEN-induced mutations as determined by the T7E1 assay is shown. Arrows indicate the expected positions of DNA bands cleaved by mismatch-sensitive T7E1. The numbers at the bottom of the gel indicate mutation percentages calculated by band intensities. (**c,f**) DNA sequences of the wild-type (WT) and mutant clones, with RGEN recognition sites shown in red and the protospacer adjacent motif (PAM) sequence in bold characters. Dashes and lower case letters indicate deleted and inserted bases, respectively (the number of inserted or deleted bases are described in the parentheses; out-of-frame and in-frame mutations are indicated using red and black colours, respectively). The number of occurrences is shown in parentheses; X1 and X2 indicate the number of each clone. Mutation frequencies were obtained by dividing the number of mutant clones by the number of total clones.

mRFP<sup>+</sup>eGFP<sup>+</sup> cells was 33%, 7.4-fold higher than that in unsorted cells (4.5%), corroborating the efficient enrichment of cells containing RGEN-induced mutations (Fig. 2c).

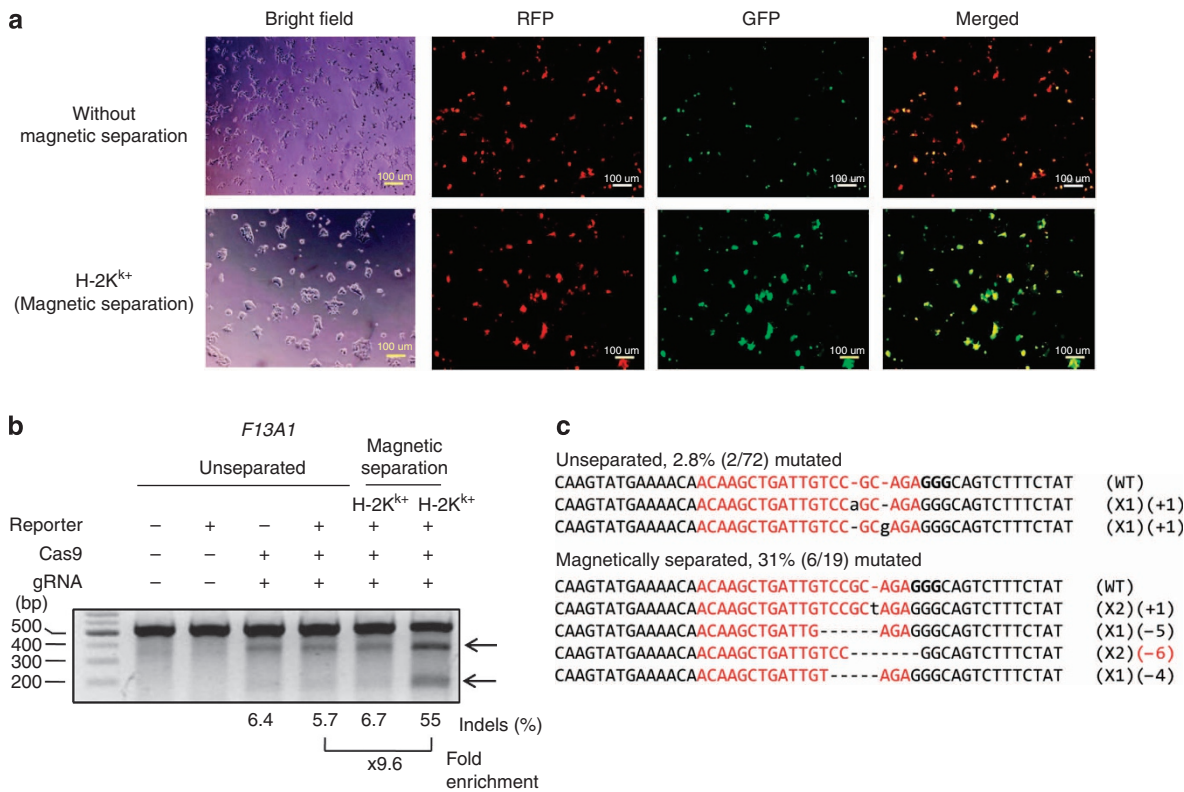
Next, to determine whether this reporter system is applicable to other RGENs, we tested both *F13A1* and *PPP1R1B*-targeting RGENs. Three days after cotransfecting *F13A1*- and *PPP1R1B*-targeting RGEN and reporter plasmids, 8.2 and 7.8% of HEK293T cells were mRFP<sup>+</sup>eGFP<sup>+</sup> cells (Fig. 3a,d). Following the isolation of genomic DNA from the sorted cells, the T7E1 assay revealed that the mutation frequencies in these cells were 36% and 53%, respectively, which is 7.8- and 4.6-fold higher than in unsorted cells (Fig. 3b,e), suggesting that reporter-based flow cytometric sorting is portable to other RGENs for the enrichment of mutant cells. This enrichment was also validated by sequencing the genomic DNA around the target site, which revealed that the mutation frequency after *F13A1* and *PPP1R1B*-targeting RGEN treatment was 10-fold (32%) and 4.3-fold (23%) higher, respectively, in mRFP<sup>+</sup>eGFP<sup>+</sup> cells than in unsorted cells (Fig. 3c,f). In line with the result with *BMP1*-targeting RGENs, the mutation frequencies of cells transfected with only Cas9 and guide RNAs without the reporter plasmid, unsorted cells and mRFP<sup>+</sup>eGFP<sup>+</sup> cells were all comparable (in the *F13A1* group, 5.0%, 4.6% and 5.2%, respectively; in the *PPP1R1B* group, 11%, 11% and 12%, respectively). We next tested this reporter system using HeLa cells, a human cervical cancer cell line. Flow cytometry showed that 6.7% of these cells were mRFP<sup>+</sup>eGFP<sup>+</sup> and the T7E1 assay revealed that the mutation frequency in the

sorted cells was 52%, 4.5-fold higher than in unsorted cells (Supplementary Fig. 2), suggesting that this flow cytometric sorting is also applicable to other cell types for the enrichment of cells containing RGEN-induced mutations.

### Magnetic separation-mediated enrichment of mutant cells.

We next tested whether magnetic separation can enrich cells containing RGEN-induced mutations. For this experiment, we used magnetic reporters<sup>27</sup>, in which both eGFP and H-2K<sup>k</sup> are expressed by frame-shifting indel formation in the target sequence of the reporter (Supplementary Fig. 3). Three days after the cotransfection of the reporter plasmid, a plasmid encoding Cas9 and a plasmid encoding *F13A1*-targeting crRNA and tracrRNA into HEK293T cells (Supplementary Fig. 4), we isolated H-2K<sup>k</sup> expressing cells by magnetic separation using a magnetic bead-conjugated anti-H2K<sup>k</sup> antibody. Twenty-four hours after these cells were plated into dishes, fluorescent microscopic observation showed that the population of magnetically separated cells was enriched with eGFP<sup>+</sup> cells (Fig. 4a), suggesting enrichment of cells containing RGEN-induced indels at the target site. The T7E1 assay revealed that the mutation frequency in the magnetically separated cells was 55%, 9.6-fold higher than that in unseparated cells (5.7%) (Fig. 4b), indicating efficient enrichment of cells containing RGEN-induced *F13A1* mutations. The mutation frequencies of cells transfected with only Cas9 and guide RNAs without the





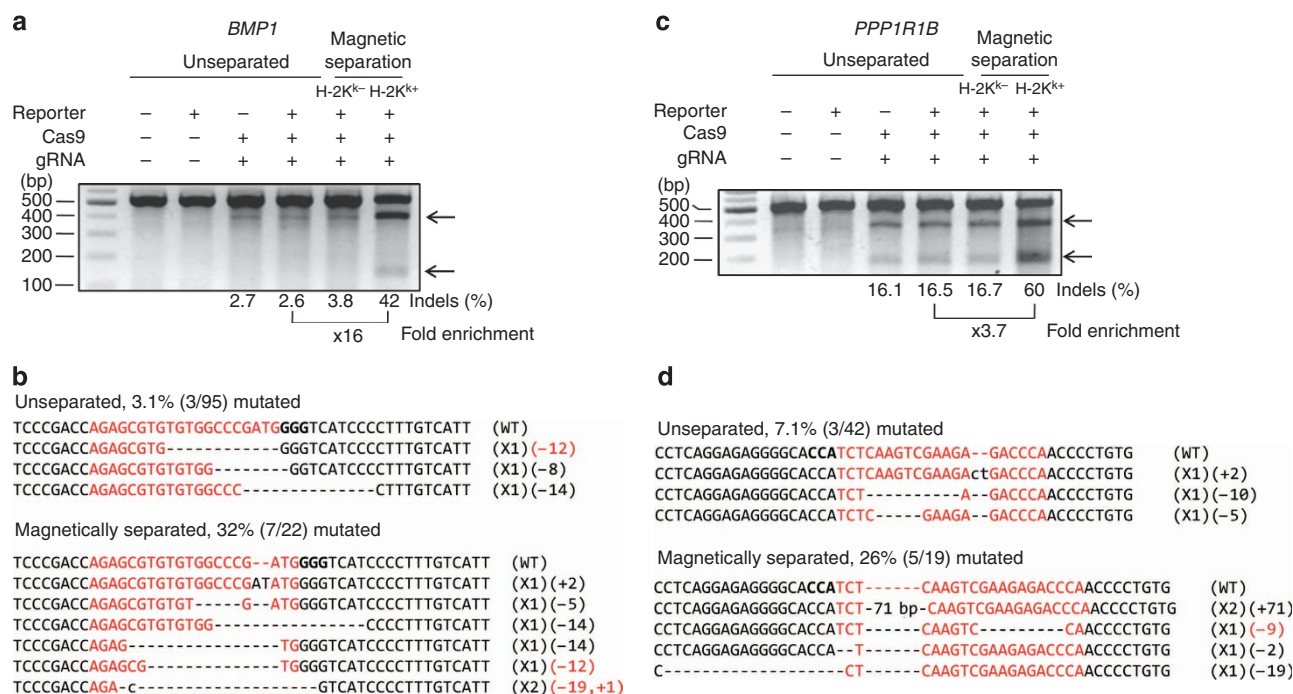
**Figure 4 | Magnetic separation-mediated enrichment of *F13A1*-disrupted cells.** (a) Fluorescent microscopy of HEK293T cells cultured for 1 day after magnetic separation, which was performed 3 days after cotransfection of the reporter plasmid, a plasmid encoding Cas9 and a plasmid encoding *F13A1*-targeting crRNA and tracrRNA. Magnetic separation enriched eGFP<sup>+</sup> cells. Scale bar, 100 μm. (b) The frequency of RGEN-induced mutations as determined by the T7E1 assay is shown. Arrows indicate the expected positions of DNA bands cleaved by mismatch-sensitive T7E1. The numbers at the bottom of the gel indicate mutation percentages calculated by band intensities. (c) DNA sequences of the wild-type (WT) and mutant clones, with RGEN recognition sites shown in red and the protospacer adjacent motif (PAM) sequence in bold characters. Dashes and lower case letters indicate deleted and inserted bases, respectively (the number of inserted or deleted bases are described in the parentheses; out-of-frame and in-frame mutations are indicated using red and black colours, respectively). The number of occurrences is shown in parentheses; X1 and X2 indicates the number of each clone. Mutation frequencies were obtained by dividing the number of mutant clones by the number of total clones.

reporter plasmid, unsorted cells and mRFP<sup>+</sup>eGFP<sup>-</sup> cells were all similar (6.4%, 5.7% and 6.7%, respectively). Sequencing the genomic DNA around the target site after PCR amplification revealed that the mutation frequency in the magnetically separated cells was 31%, 11-fold higher than in the unseparated cells (2.8%) (Fig. 4c), corroborating the drastic enrichment of mutant cells mediated by magnetic separation.

When we tested this magnetic reporter using *BMP1*- and *PPP1R1B*-targeting RGENs, the T7E1 assay showed that the mutation frequency was 42% and 60% in magnetically separated cells, respectively, which is 16- and 3.7-fold higher than in unseparated cells (2.6% and 16.5%, respectively) (Fig. 5a,c). This enrichment was confirmed by sequencing the target region, which revealed that the mutation frequency after *BMP1*- and *PPP1R1B*-targeting RGEN treatment was 10-fold (32%) and 3.7-fold (26%) higher, respectively, in magnetically separated cells than in unseparated cells (Fig. 5b,d). This result indicates that cells in which mutations are induced by other RGENs can be efficiently enriched using magnetic reporters. Compatible with the result using *F13A1*-targeting RGENs, the mutation frequencies of cells transfected with only Cas9 and guide RNAs without the reporter plasmid, unsorted cells and mRFP<sup>+</sup>eGFP<sup>-</sup> cells were all comparable (in the *BMP1* group, 2.7%, 2.6% and 3.8%, respectively; in the *PPP1R1B* group, 16.1%, 16.5% and 16.7%, respectively). This magnetic separation-mediated enrichment of mutant cells was also validated in other cell types, as evidenced by

3.8-fold enrichment of HeLa cells containing RGEN-induced mutations, reaching a mutation frequency of 36% as calculated by the T7E1 assay (Supplementary Fig. 5).

**Mutant cell enrichment using hygromycin selection.** The separation methods discussed above require instruments such as flow cytometers and magnetic separators, and both necessitate trypsinization of adherent cells to make single-cell suspensions. However, antibiotic selection does not require any special instruments or antibodies. Thus, we attempted to enrich cells containing RGEN-induced mutations using antibiotic selection. For this approach, we used a hygromycin reporter<sup>27</sup> that expresses a hygromycin-resistance protein (Hygro<sup>R</sup>)-eGFP fusion protein only when frame-shifting indels are generated in the target sequences of the reporter by nuclease activity (Supplementary Figs 6 and 7). Hygromycin treatment for 2 days starting at 2 days after the cotransfection of the reporter plasmid, a plasmid encoding Cas9 and a plasmid encoding *PPP1R1B*-targeting crRNA and tracrRNA resulted in the death of most cells; the remaining population of cells was enriched for eGFP<sup>+</sup> cells (Fig. 6a). The T7E1 assay showed that the mutation frequency at the *PPP1R1B* gene in the hygromycin-selected cells was 57%, 8.6-fold higher than that in unselected cells (6.6%) (Fig. 6b), indicating efficient enrichment of mutant cells. DNA sequencing of the region containing the target site confirmed this



**Figure 5 | Magnetic separation-mediated enrichment of *BMP1*- and *PPP1R1B*-disrupted cells.** Magnetic separation was performed 3 days after cotransfection of the reporter plasmid, a plasmid encoding Cas9 and a plasmid encoding *BMP1*- (**a,b**) or *PPP1R1B*- (**c,d**) targeting crRNA and tracrRNA into HEK293T cells. (**a,c**) The frequency of RGEN-induced mutations was determined by the T7E1 assay. Arrows indicate the expected positions of DNA bands cleaved by mismatch-sensitive T7E1. The numbers at the bottom of the gel indicate mutation percentages calculated by band intensities. (**b,d**) DNA sequences of the wild-type (WT) and mutant clones, with RGEN recognition sites shown in red and the protospacer adjacent motif (PAM) sequence in bold characters. Dashes and lower case letters indicate deleted and inserted bases, respectively (the number of inserted or deleted bases are described in the parentheses; X1 and X2 indicate the number of each clone. Mutation frequencies were obtained by dividing the number of mutant clones by the number of total clones.

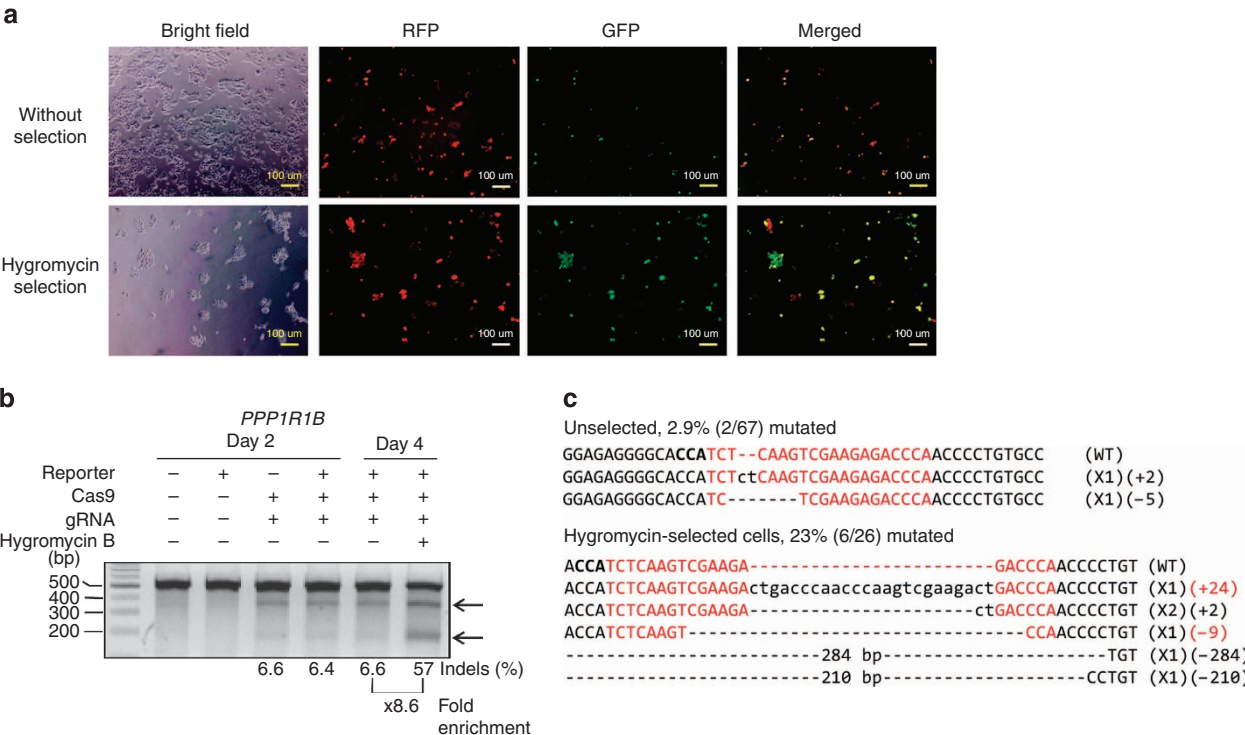
enrichment by revealing that the mutation frequency in the hygromycin-resistant cells was 23%, 7.7-fold higher than that in unselected cells (Fig. 6c).

When we tested this enrichment approach using *BMP1*- and *F13A1*-targeting RGENs, we observed mutation frequencies of 33 and 34%, respectively, after hygromycin selection, representing 12-fold and 5.6-fold enrichment of mutant cells (unselected cells displayed mutation frequencies of 2.8% and 6.1%, respectively) (Fig. 7a,c). This enrichment was corroborated by sequencing the genomic DNA around the target site, which revealed that the mutation frequency after transfection of plasmids encoding *BMP1* and *F13A1*-targeting RGENs was 6.7-fold (20%) and 7.3-fold (25%) higher, respectively, in the selected cells than in unselected cells (Fig. 7b,d). The mutation frequencies of cells transfected with only Cas9 and guide RNAs without the reporter plasmid, unselected cells at day 2 and unselected cells at day 4 after transfection were all comparable (in the *PPP1R1B* group, 6.6%, 6.4% and 6.6%, respectively; in the *BMP1* group, 3.0%, 2.5% and 2.8%, respectively; in the *F13A1* group, 6.9%, 5.8% and 6.1%, respectively), which are compatible with the observations using flow cytometric and magnetic reporters. Furthermore, hygromycin selection of HeLa cells revealed a 5.4-fold enrichment of cells containing RGEN-induced mutations (Supplementary Fig. 8), indicating that the hygromycin system, like the flow cytometric and magnetic systems, is also applicable to other cell types.

**Comparison of reporters and enrichment methods.** In summary, the three different enrichment methods, each tested using three RGENs, enriched mutant cells from 3.7–11-fold

(mean 7.7-fold, median 7.4-fold) (Table 1). Interestingly, the fold enrichment tends to be high when RGEN activities are low (Fig. 8a), leading to comparable mutation frequencies that ranged from 20 to 33% (fold difference between the highest/lowest = 1.7) after the selection; in comparison, the mutation frequencies before enrichment ranged from 2.8 to 7.1% (fold difference between the highest/lowest = 2.6) (Fig. 8b). In our current experiments, the fold enrichments were similar among the three selection methods (Fig. 8c), which is also in line with previous studies using ZFNs and TALENs<sup>27</sup>. Thus, researchers can choose the reporter system that best fits their experimental conditions and research environment on the basis of previously described pros and cons of each reporter<sup>27</sup>.

**Reporter cotransfection effects.** Our previous study using ZFNs showed that reporter-mediated selection results in the enrichment of highly transfected cells, which partially accounts for the enrichment of mutant cells<sup>24</sup>. To determine whether this phenomenon is true for RGENs, we sorted cells depending on RFP expression (an indicator of the level of transfection) and evaluated the mutation frequencies (Supplementary Fig. 9). As expected, the T7E1 assay revealed that cells containing mutations were enriched in the following order: RFP<sup>bright</sup> (31% with mutations), RFP<sup>medium</sup> (24%) and RFP<sup>dim</sup> (9.2%), indicating that high levels of transfection can lead to high mutation frequencies. However, the degree of enrichment of gene-disrupted cells even in the RFP<sup>bright</sup> population was lower than that in the RFP<sup>+</sup>GFP<sup>+</sup> population (48%), suggesting that our reporter system is more useful for enrichment than simple sorting based



**Figure 6 | Hygromycin selection-mediated enrichment of *PPP1R1B*-disrupted cells. (a)** Fluorescent microscopy of HEK293T cells treated with hygromycin for 2 days starting 2 days after the cotransfection of the reporter plasmid, a plasmid encoding Cas9 and a plasmid encoding *PPP1R1B*-targeting crRNA and tracrRNA. Hygromycin selection-enriched eGFP<sup>+</sup> cells. Scale bar, 100  $\mu$ m. **(b)** The frequency of RGEN-induced mutations as determined by the T7E1 assay. Arrows indicate the expected positions of DNA bands cleaved by mismatch-sensitive T7E1. The numbers at the bottom of the gel indicate mutation percentages calculated by band intensities. **(c)** DNA sequences of the wild-type (WT) and mutant clones, with RGEN recognition sites shown in red and the protospacer adjacent motif (PAM) sequence in bold characters. Dashes and lower case letters indicate deleted and inserted bases, respectively (the number of inserted or deleted bases are described in the parentheses; out-of-frame and in-frame mutations are indicated using red and black colours, respectively). The number of occurrences is shown in parentheses; X1 and X2 indicates the number of each clone. Mutation frequencies were obtained by dividing the number of mutant clones by the number of total clones.

on transfection efficiency alone, which is in line with our previous results using ZFNs<sup>24</sup>.

We next determined whether the target sequence in the reporter plasmid competes with that in the genomic DNA, thereby attenuating the RGEN-induced editing of the genomic target sequence. We cotransfected various amounts of a reporter plasmid containing the target sequence together with a fixed amount of plasmid encoding an RGEN (targeting *PPP1R1B*). To make the total amount of transfected DNA constant, we cotransfected compensating amounts of a reporter plasmid with a different target sequence (from the *BMP1* gene in this case) from that of the RGEN. The T7E1 assay revealed that the efficiency of RGEN-induced gene disruption at the target sequence on the host chromosome was not changed by the transfection of a high amount of target sequence-bearing reporter plasmid (Supplementary Fig. 10), suggesting that the presence of many copies of target sequence in the transfected reporter plasmid does not attenuate the creation of RGEN-induced gene mutations at the target sequence on the host chromosome. This conclusion is in line with our previous results showing that cotransfection of a target sequence-containing reporter plasmid does not affect the efficiency of ZFN- and TALEN-directed genome editing<sup>24,27</sup>.

**Out-of-frame mutation ratio and off-target effects.** We also determined the frequency of null (out-of-frame) versus in-frame mutations before and after enrichment. The frequency of out-of-frame mutations among total mutations ranged from 57 to 100%

and was not changed significantly in nine independent enrichments (Supplementary Table 1). The frequency of out-of-frame mutations that cause gene disruption among the total sequenced alleles ranged from 2.1–7.1% before enrichment and from 13–33% after enrichment, resulting in 2.6–11-fold enrichment of out-of-frame mutations, which is similar to enrichment folds of total mutations.

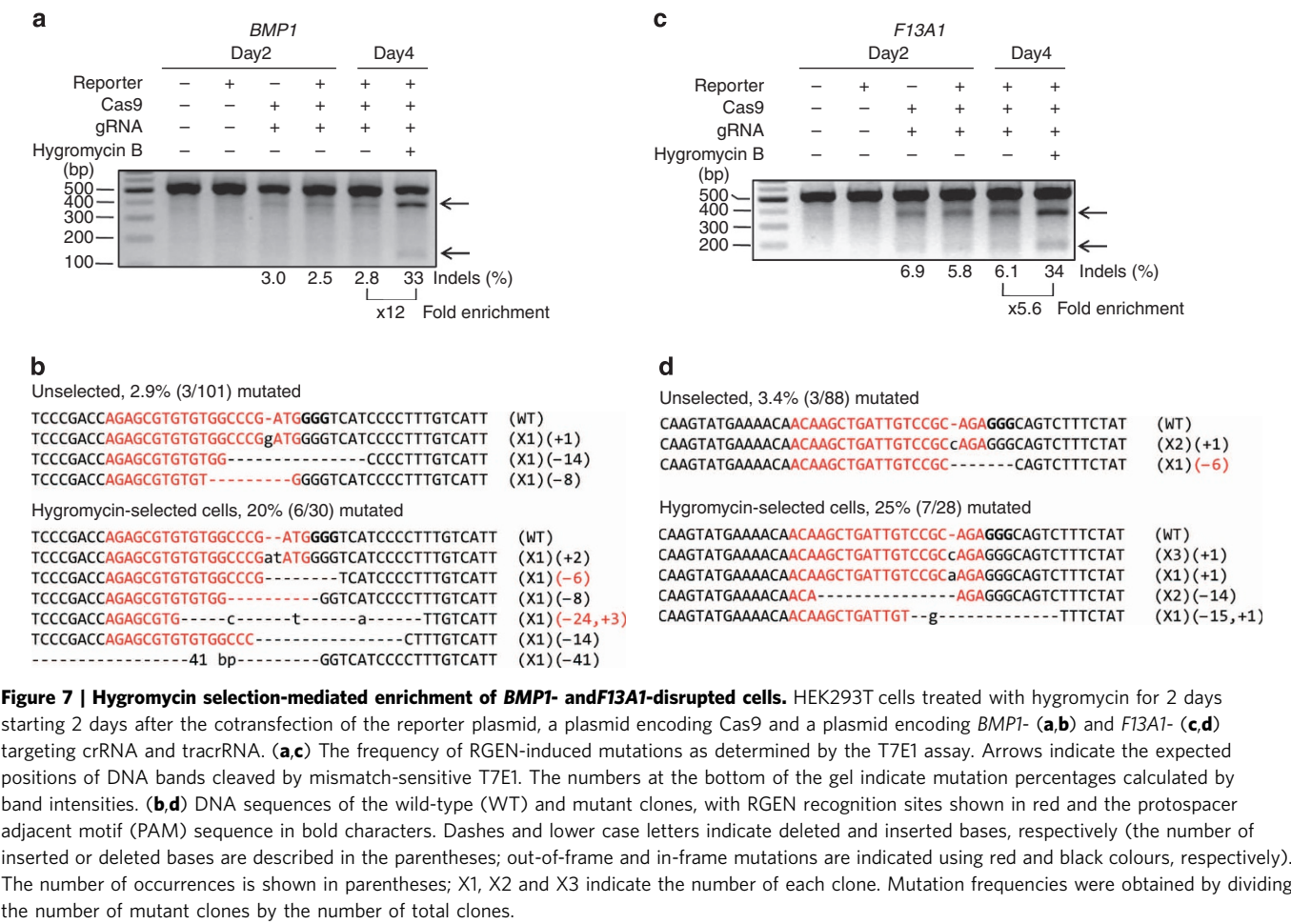
Recently, it has been reported that RGENs can have off-target effects<sup>29–35</sup>. When we determined the mutation frequencies at the four to five top off-target sites of all three RGENs used in the study, we found that we did not detect off-target effects even after enrichment (Supplementary Figs 11–13), raising the possibility that our enrichment process does not exacerbate off-target effects.

**Discussion**

Here we showed that cells containing RGEN-induced mutations can be enriched using flow cytometry, magnetic separation or hygromycin selection. Given that the fold enrichment tends to be higher when RGENs with low activity are used, our enrichment methods will be particularly useful when RGENs with high activity are not available; thus, these reporters complement RGENs so that RGENs coupled with reporters represent a reliable approach for making gene-disrupted cells.

Partly because RGENs are easier and simpler to prepare in-house than ZFNs or TALENs, RGENs are expected to be used more widely as a genome-editing tool and to be preferred by beginners in genome editing. Such beginners would be more likely to design RGENs with low activity and/or to make mistakes





**Figure 7 | Hygromycin selection-mediated enrichment of *BMP1*- and *F13A1*-disrupted cells.** HEK293T cells treated with hygromycin for 2 days starting 2 days after the cotransfection of the reporter plasmid, a plasmid encoding Cas9 and a plasmid encoding *BMP1*- (**a,b**) and *F13A1*- (**c,d**) targeting crRNA and tracrRNA. (**a,c**) The frequency of RGEN-induced mutations as determined by the T7E1 assay. Arrows indicate the expected positions of DNA bands cleaved by mismatch-sensitive T7E1. The numbers at the bottom of the gel indicate mutation percentages calculated by band intensities. (**b,d**) DNA sequences of the wild-type (WT) and mutant clones, with RGEN recognition sites shown in red and the protospacer adjacent motif (PAM) sequence in bold characters. Dashes and lower case letters indicate deleted and inserted bases, respectively (the number of inserted or deleted bases are described in the parentheses; out-of-frame and in-frame mutations are indicated using red and black colours, respectively). The number of occurrences is shown in parentheses; X1, X2 and X3 indicate the number of each clone. Mutation frequencies were obtained by dividing the number of mutant clones by the number of total clones.

| Table 1   Efficiencies of mutant cell enrichment using surrogate reporters. |                        |                  |                 |                      |
|---|------------------------|------------------|-----------------|----------------------|
| Target gene   | Mutation frequency (%) |                  | Fold enrichment | Enrichment method    |
|   | Before enrichment      | After enrichment |                 |                      |
| <i>BMP1</i>   | 4.5 (3/67)             | 33 (7/21)        | 7.4             | Flow cytometry       |
|   | 3.2 (3/95)             | 32 (7/22)        | 10              | Magnetic separation  |
|   | 3.0 (3/101)            | 20 (6/30)        | 6.7             | Hygromycin selection |
| <i>F13A1</i>  | 3.1 (12/387)           | 32 (12/37)       | 10              | Flow cytometry       |
|   | 2.8 (2/72)             | 32 (6/19)        | 11              | Magnetic separation  |
|   | 3.4 (3/88)             | 25 (7/28)        | 7.3             | Hygromycin selection |
| <i>PPP1R1B</i>  | 5.3 (3/57)             | 23 (5/22)        | 4.3             | Flow cytometry       |
|   | 7.1 (3/42)             | 26 (5/19)        | 3.7             | Magnetic separation  |
|   | 3.0 (2/67)             | 23 (6/26)        | 7.7             | Hygromycin selection |

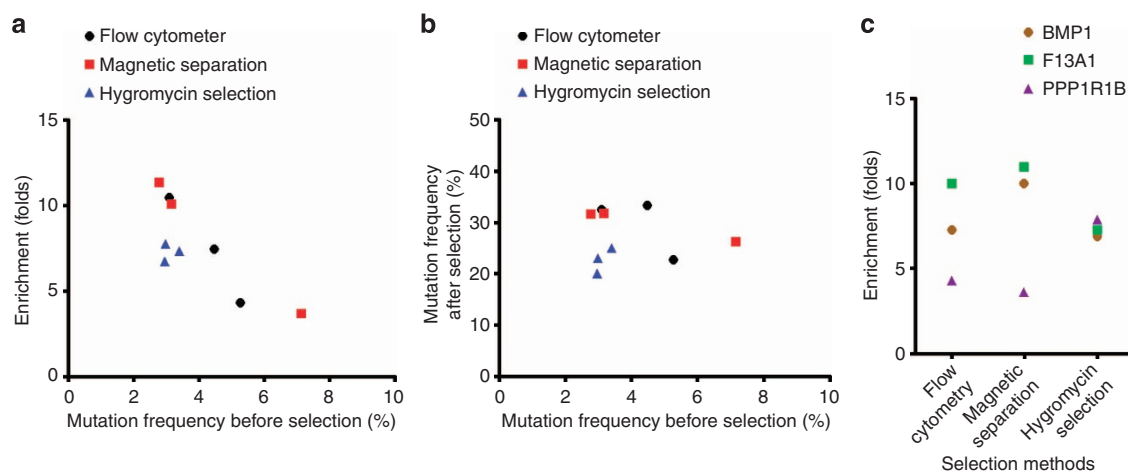
The mutation frequency was calculated by sequencing the genomic DNA containing the RGEN-target sites after PCR amplification. The ratios in the parentheses represent the number of mutation sequences/the total number of sequenced amplicons.

in one or more genome-editing steps. Our reporters will be useful in this situation because they enable both monitoring of RGEN activity in live cells with fluorescent microscopy and the enrichment of genome-edited cells, a process that is more robust with low activity RGENs. Furthermore, in all three reporter systems, the enrichment process can also be monitored indirectly with fluorescent microscopy by evaluating the fraction of GFP<sup>+</sup> cells.

We found that the addition of an episomal target sequence did not attenuate the RGEN-induced editing of the genomic

target sequence. We tested a range of amounts of the reporter-incorporated episomal target sequence, up to ninefold higher than the ordinary amount of reporter (we could not evaluate the effect of adding even more reporters because the amount of DNA that can be transfected is limited). This lack of effect on RGEN activity suggests that the amount of guide RNA and Cas9 protein is sufficient to edit both the episomal and genomic target sequences within this tested range. This unaltered RGEN activity at the genomic target site is compatible with previous results<sup>24,27</sup> and is one of the practical advantages of our reporter system.





**Figure 8 | Factors affecting fold enrichment.** (a,b) The association of fold enrichment, mutation frequencies before selection and mutation frequencies after selection are shown. Mutation frequencies were calculated from sequencing results. Each dot represents an independent experiment. The selection methods used are symbolized with different coloured symbols. (c) The relationship between fold enrichment and selection method. Each dot represents an independent experiment. The targeted genes are symbolized using different coloured symbols.

RGENs are emerging as an approachable genome-editing tool. This is the first study showing that RGEN-induced mutations can be enriched using surrogate reporters. The three reporters resulted in comparable enrichment; researchers can choose a reporter system on the basis of their experimental conditions. We envision that these three types of reporters will facilitate the use of RGENs in a wide range of biomedical studies.

## Methods

**Reporters.** Reporters were prepared as previously described<sup>24,27</sup>. Briefly, oligonucleotides including target sequences (Supplementary Table 2) were synthesized (Bioneer, Daejeon, South Korea) and annealed *in vitro* using a thermocycler (95 °C for 5 min and then ramped down to 25 °C at 5 °C per min). The annealed oligonucleotides were ligated into the reporter vectors digested with EcoRI and BamHI.

**RGENs.** The Cas9-expressing p3s-Cas9HC plasmid has been previously described<sup>4</sup> and is available from Addgene (#43945). To construct the pRG2-CT plasmid, which expresses crRNA and tracrRNA under the control of the hU6 promoter, U6-crRNA and U6-tracrRNA cassettes were generated in the pLKO.1-puro plasmid and then sequentially subcloned into the minimal pUC18 backbone plasmid. The full sequence of this vector is described in Supplementary Fig. 14 and the vector map is shown in Supplementary Fig. 15. Oligonucleotides containing each target sequences were synthesized (Macrogen, Seoul, South Korea) and annealed *in vitro* using a thermocycler. The pRG2-CT vector was digested with BsaI and ligated with annealed oligonucleotides. Oligonucleotide sequences are listed in Supplementary Table 3.

**Cell culture.** Human embryonic kidney 293T (HEK293T) cells and HeLa cells (human cervical cancer cells) were purchased from American Type Culture Collection (Manassas, VA) and maintained in Dulbecco's modified Eagle medium (DMEM; Invitrogen, Carlsbad, CA, USA) supplemented with 100 units ml<sup>-1</sup> penicillin, 100 µg ml<sup>-1</sup> streptomycin and 10% fetal bovine serum.

**Transfection.** Cells were transfected using lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA) or polyethyleneimine (linear, MW ~ 25,000; Polysciences, Warrington, PA, USA) at a weight ratio of 1:2:1 (plasmid encoding a Cas9: plasmid encoding crRNA and tracrRNA: reporter) unless specified otherwise.

**Flow cytometry.** Adherent cells were trypsinized and resuspended in 2% FBS in PBS. Single-cell suspensions were analysed and sorted using the FACSaria II cell sorter (BD Biosciences, San Jose, CA, USA). To collect cells that contain RGEN-induced mutations, cells with strong eGFP signals were sorted. Untransfected cells and cells transfected with reporters alone were used as controls.

**Magnetic separation.** Magnetic separations were performed as previously described<sup>27</sup>. Briefly, the transfected cells were cultured for 3 days and trypsinized. Cell suspensions were mixed with a magnetic bead-conjugated antibody against

H-2K<sup>k</sup> (MACSelect K<sup>k</sup> microbeads; Miltenyi Biotech, Germany) at 80 µl antibody per 1 × 10<sup>7</sup> cells and incubated for 20 min at 4 °C. Labelled cells were separated using a column (MACS LS column; Miltenyi Biotech) according to the manufacturer's instructions.

**Hygromycin selection.** Two days after transfection, hygromycin (Invitrogen, Carlsbad, CA, USA) was added to the culture media at a final concentration of 2 mg ml<sup>-1</sup> for 2 days.

**T7E1 assay.** The T7E1 assay was performed as previously described<sup>24,26,36</sup>. Briefly, after the isolation of genomic DNA using the DNeasy Blood & Tissue Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions, the region of DNA containing the nuclease target site was nested PCR-amplified using appropriate primers (Supplementary Table 4 and 5). The amplicons were denatured by heating and annealed to form heteroduplex DNA, which was treated with 5 units of T7 endonuclease 1 (New England Biolabs) for 15–20 min at 37 °C and analysed using 2% agarose gel electrophoresis. Mutation frequencies were calculated as previously described based on the band intensities using Image J software and the following equation<sup>37</sup>: mutation frequency (%) = 100 × (1 – (1 – fraction cleaved)<sup>1/2</sup>), where the fraction cleaved is the total relative density of the cleavage bands divided by the sum of the relative density of the cleavage bands and uncut bands.

**Sequencing analysis.** Sequencing of the region including the target sequence was performed as previously described<sup>24,27</sup>. Briefly, PCR amplicons that included RGEN-target sites were cloned into the T-Blunt vector. Cloned plasmids were sequenced using the primers used for PCR amplification.

## References

- Horvath, P. & Barrangou, R. CRISPR/Cas, the immune system of bacteria and archaea. *Science* **327**, 167–170 (2010).
- Wiedenheft, B., Sternberg, S. H. & Doudna, J. A. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **482**, 331–338 (2012).
- Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
- Cho, S. W., Kim, S., Kim, J. M. & Kim, J. S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* **31**, 230–232 (2013).
- Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
- Friedland, A. E. *et al.* Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. *Nat. Methods* **10**, 741–743 (2013).
- Gratz, S. J. *et al.* Genome engineering of drosophila with the CRISPR RNA-guided Cas9 nuclease. *Genetics* **194**, 1029–1035 (2013).
- Hwang, W. Y. *et al.* Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 227–229 (2013).
- Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L. A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* **31**, 233–239 (2013).

10. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
11. Li, J. F. *et al.* Multiplex and homologous recombination-mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. *Nat. Biotechnol.* **31**, 688–691 (2013).
12. Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J. D. & Kamoun, S. Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* **31**, 691–693 (2013).
13. Shan, Q. *et al.* Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 686–688 (2013).
14. Li, W., Teng, F., Li, T. & Zhou, Q. Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems. *Nat. Biotechnol.* **31**, 684–686 (2013).
15. Li, D. *et al.* Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 681–683 (2013).
16. Wang, H. *et al.* One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910–918 (2013).
17. Jinek, M. *et al.* RNA-programmed genome editing in human cells. *eLife* **2**, e00471 (2013).
18. Ding, Q. *et al.* Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. *Cell Stem Cell* **12**, 393–394 (2013).
19. Yang, L. *et al.* Optimization of scarless human stem cell genome editing. *Nucleic Acids Res.* **41**, 9049–9061 (2013).
20. Hauschild, J. *et al.* Efficient generation of a biallelic knockout in pigs using zinc-finger nucleases. *Proc. Natl Acad. Sci. USA* **108**, 12013–12017 (2011).
21. Ding, Q. *et al.* A TALEN genome-editing system for generating human stem cell-based disease models. *Cell Stem Cell* **12**, 238–251 (2013).
22. Frank, S., Skryabin, B. V. & Greber, B. A modified TALEN-based system for robust generation of knock-out human pluripotent stem cell lines and disease models. *BMC Genomics* **14**, 773 (2013).
23. Carlson, D. F. *et al.* Efficient TALEN-mediated gene knockout in livestock. *Proc. Natl Acad. Sci. USA* **109**, 17382–17387 (2012).
24. Kim, H., Um, E., Cho, S. R., Jung, C. & Kim, J. S. Surrogate reporters for enrichment of cells with nuclease-induced mutations. *Nat. Methods* **8**, 941–943 (2011).
25. Perez, E. E. *et al.* Establishment of HIV-1 resistance in CD4<sup>+</sup> T cells by genome editing using zinc-finger nucleases. *Nat. Biotechnol.* **26**, 808–816 (2008).
26. Kim, H. J., Lee, H. J., Kim, H., Cho, S. W. & Kim, J. S. Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Res.* **19**, 1279–1288 (2009).
27. Kim, H., Kim, M. S., Wee, G., Lee, C. I. & Kim, J. S. Magnetic separation and antibiotics selection enable enrichment of cells with ZFN/TALEN-induced mutations. *PLoS One* **8**, e56476 (2013).
28. Lee, H. J., Kim, E. & Kim, J. S. Targeted chromosomal deletions in human cells using zinc finger nucleases. *Genome Res.* **20**, 81–89 (2010).
29. Cho, S. W. *et al.* Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.* **24**, 132–141 (2013).
30. Cradick, T. J., Fine, E. J., Antico, C. J. & Bao, G. CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res.* **41**, 9584–9592 (2013).
31. Hruscha, A. *et al.* Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. *Development* **140**, 4982–4987 (2013).
32. Mali, P. *et al.* CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* **31**, 833–838 (2013).
33. Pattanayak, V. *et al.* High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat. Biotechnol.* **31**, 839–843 (2013).
34. Fu, Y. *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* **31**, 822–826 (2013).
35. Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832 (2013).
36. Ramakrishna, S., Kim, Y. H. & Kim, H. Stability of zinc finger nuclease protein is enhanced by the proteasome inhibitor MG132. *PLoS One* **8**, e54282 (2013).
37. Guschin, D. Y. *et al.* A rapid and general assay for monitoring endogenous gene modification. *Methods Mol. Biol.* **649**, 247–256 (2010).

## Acknowledgements

We are grateful to H.K.'s lab members for their suggestions and discussions on this project. H.K. and J.-S.K. are supported in part by the National Research Foundation (H.K., 2011-0013568, 2008-0062287, 2011-0019357, 2013M3A9B4076544; J.-S.K., 2013-000718) and Converging Research Center Program (H.K., 2013K000275) funded by the Ministry of Science, ICT, and Future Planning, and the Korean Health Technology R&D Project, Ministry of Health and Welfare, Republic of Korea (H.K., H10C1740).

## Author contributions

Conceived and designed the experiments: S.R., H.K. and J.-S.K.; performed the experiments: S.R., S.W.C., S.K., S.M. and R.G.; analysed the data: S.R. and H.K.; wrote the paper: H.K., S.R. and J.-S.K.

## Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Ramakrishna, S. *et al.* Surrogate reporter-based enrichment of cells containing RNA-guided Cas9 nuclease-induced mutations. *Nat. Commun.* 5:3378 doi: 10.1038/ncomms4378 (2014).